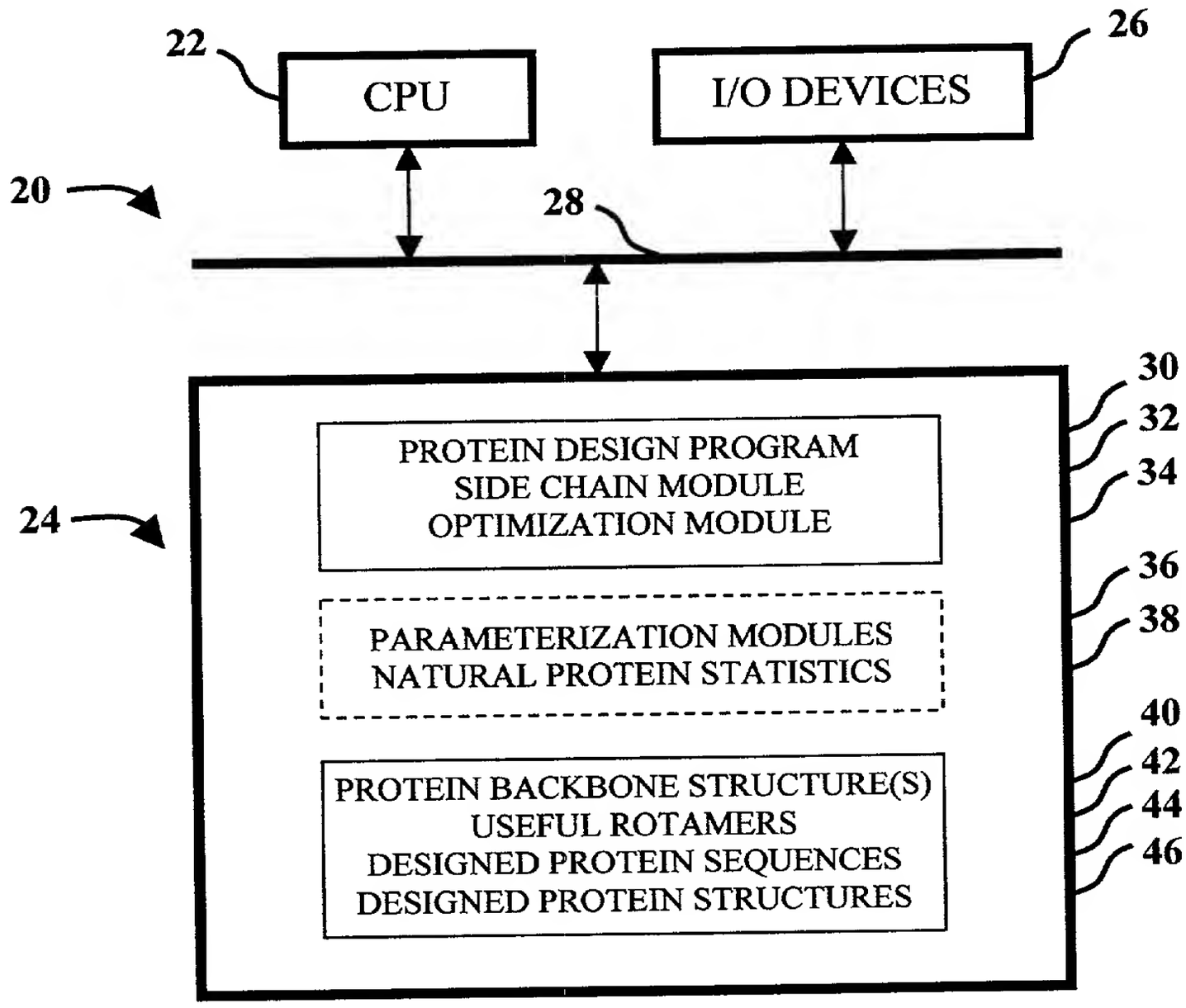
22
CPU
26
I/O DEVICES
20
28
30
32
34
24
PROTEIN DESIGN PROGRAM
SIDE CHAIN MODULE
OPTIMIZATION MODULE
36
PARAMETERIZATION MODULES
NATURAL PROTEIN STATISTICS
38
40
PROTEIN BACKBONE STRUCTURE(S)
USEFUL ROTAMERS
DESIGNED PROTEIN SEQUENCES
DESIGNED PROTEIN STRUCTURES
42
44
46

FIG. 1.

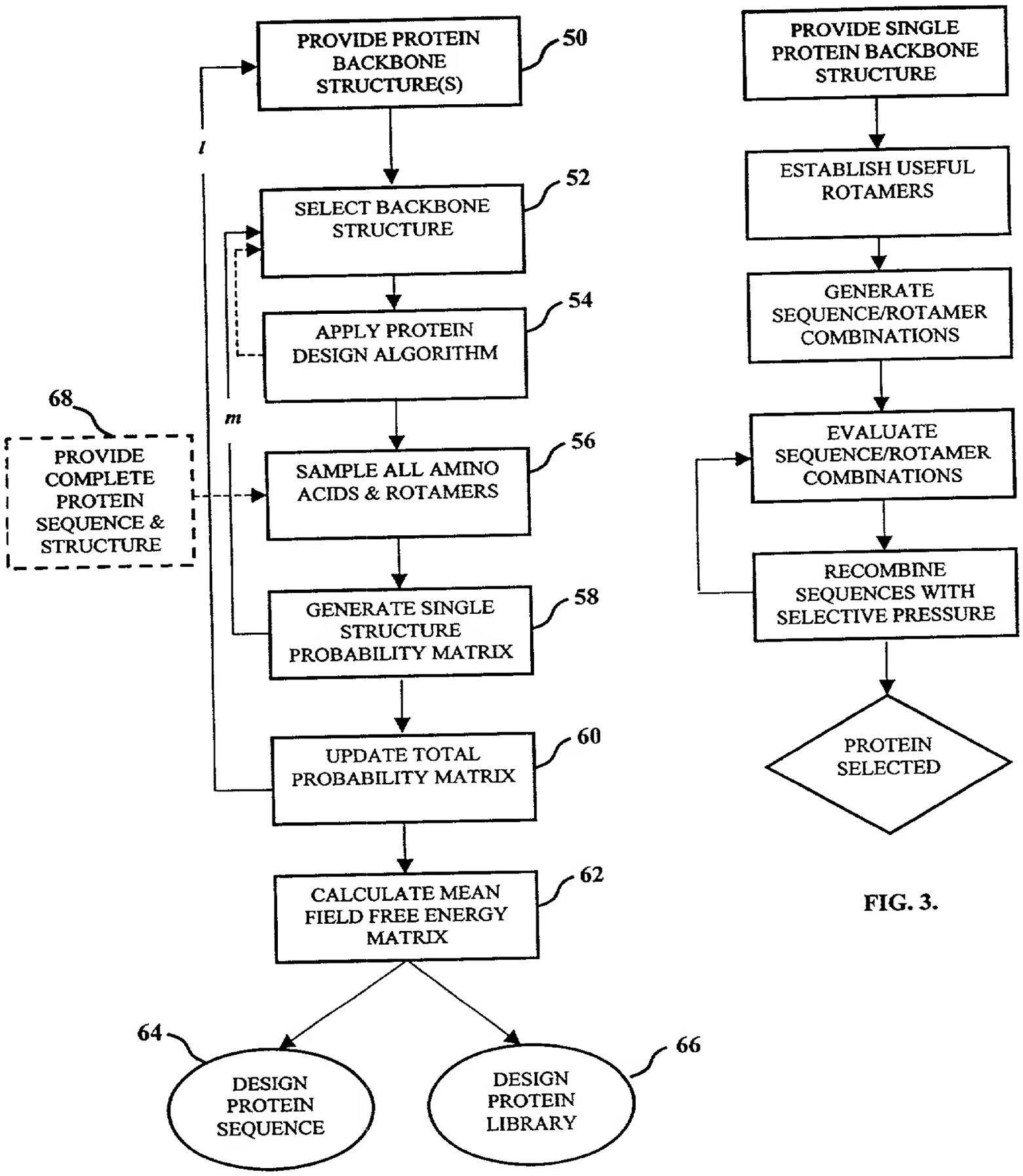
PROVIDE PROTEIN BACKBONE STRUCTURE(S)
50
SELECT BACKBONE STRUCTURE
52
APPLY PROTEIN DESIGN ALGORITHM
54
m
68
PROVIDE COMPLETE PROTEIN SEQUENCE & STRUCTURE
SAMPLE ALL AMINO ACIDS & ROTAMERS
56
GENERATE SINGLE STRUCTURE PROBABILITY MATRIX
58
UPDATE TOTAL PROBABILITY MATRIX
60
CALCULATE MEAN FIELD FREE ENERGY MATRIX
62
64
DESIGN PROTEIN SEQUENCE
66
DESIGN PROTEIN LIBRARY

FIG. 2.

PROVIDE SINGLE PROTEIN BACKBONE STRUCTURE
ESTABLISH USEFUL ROTAMERS
GENERATE SEQUENCE/ROTAMER COMBINATIONS
EVALUATE SEQUENCE/ROTAMER COMBINATIONS
RECOMBINE SEQUENCES WITH SELECTIVE PRESSURE
PROTEIN SELECTED

FIG. 3.

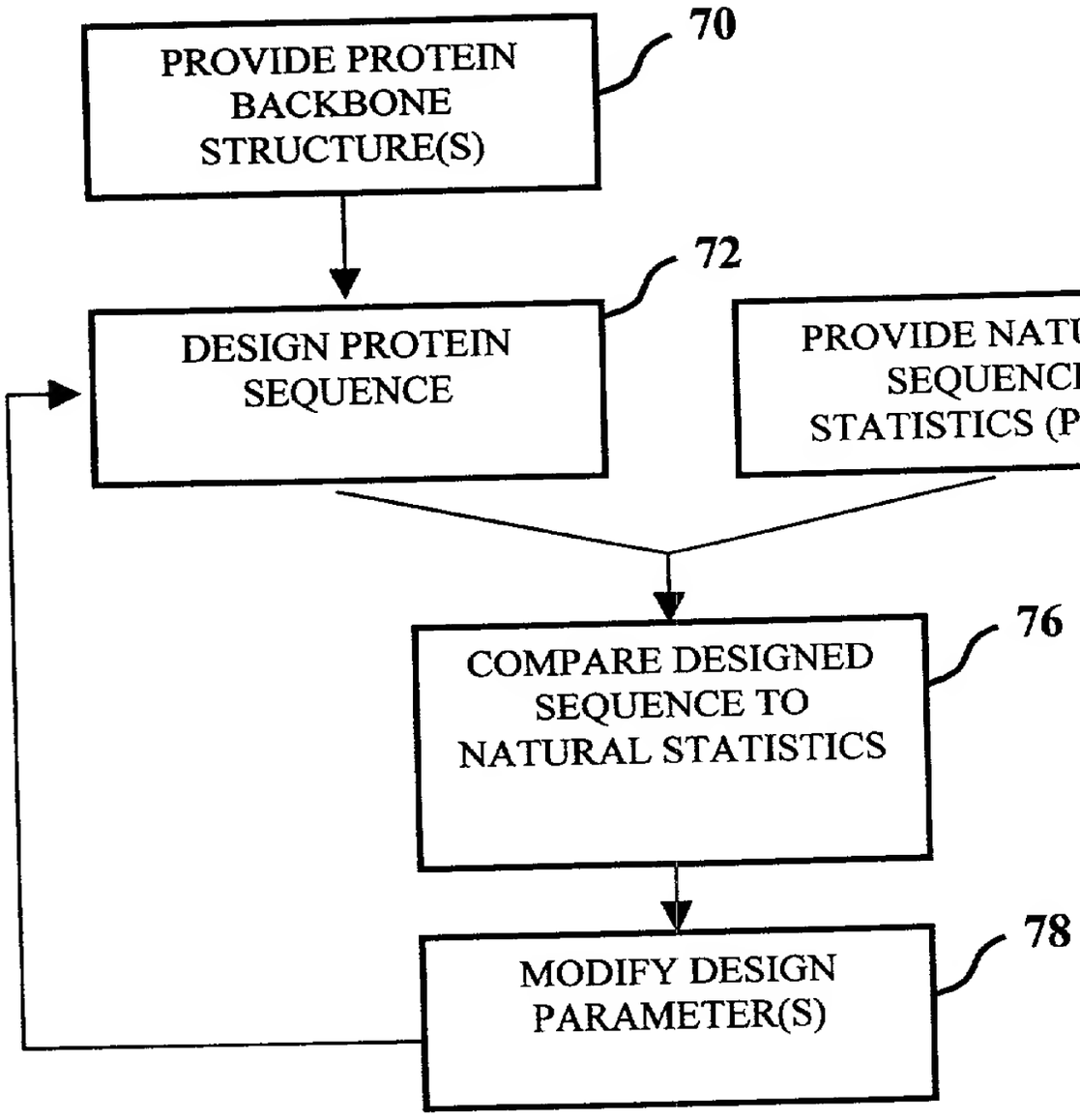
PROVIDE PROTEIN BACKBONE STRUCTURE(S)
70
DESIGN PROTEIN SEQUENCE
72
PROVIDE NATURAL SEQUENCE STATISTICS (PSSM)
74
COMPARE DESIGNED SEQUENCE TO NATURAL STATISTICS
76
MODIFY DESIGN PARAMETER(S)
78

FIG. 4.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
A	1.7	10	4.9	2	16	17	6.5	4.9	4.8	4.2	3.8	0.5	1.9	4.7	1.7	3.5	4.6	18	8.7	10	7.7	3.6	3.5	4.8	3.9	4.1	3.8	9.1	2.1	3.9	6	10	5.3	4.7
C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D	0.7	17	7.1	4	20	50	5.8	3.7	3.5	9.2	8.2	2.4	3.5	0.3	0.4	4.7	5.4	20	20	19	6.2	5.8	2.4	5.4	0.7	2.9	7.5	4.6	1.2	4.4	2.3	17	7.7	0.2
E	1.5	16	6.5	2.6	20	20	3.7	2.3	2.3	5.1	4.7	4.2	3.2	4.2	1	3.5	3.5	19	18	6.8	13	1.7	2	5.5	3.2	2.2	3.3	8.8	2.3	3	4.4	17	5.7	1.9
F	50	9.5	20	50	20	3.8	17	7.1	7.2	8	4.1	50	50	20	50	3.9	7	2	2.3	3.1	20	5.8	14	20	50	5	11	17	50	20	9.3	16	7.9	50
G	3.9	4.1	7.7	4.4	0	20	12	8.9	13	8.9	11	2.8	0.9	8.1	2.8	9.3	6.8	20	18	19	19	8.7	8.2	11	3.1	10	6.1	14	1.4	6	9	19	10	7.4
H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I	50	8.6	7.8	15	20	11	5.4	4	8.7	4.9	6.2	50	50	19	50	4.1	5.1	16	11	11	17	4.9	3	10	50	2.3	2.1	7.3	6.6	6.7	5.8	14	4.8	50
K	4.8	6.7	7.1	6	17	18	6.2	2	2.8	2	0.5	7.5	6.1	9.3	4.6	4	5.5	15	10	7.9	5.2	0.6	3	5.9	3.5	3.1	1.3	4.7	3	4.3	2.6	14	5.8	6.5
L	50	0.3	9.5	20	19	9.2	4	4.7	1.1	4.3	3.4	50	50	9.6	50	5.5	7.3	11	5.1	7.9	8.4	3.4	5.2	7.5	50	5.4	4.3	7.8	50	8.4	4.6	17	5.3	50
M	50	3.1	11	14	19	11	7	7.9	5.9	6.1	6.9	50	50	15	50	6	6.8	12	4.9	9.5	6.1	6.3	7.4	12	50	8.5	5.5	6.6	6.7	11	6.4	16	7.1	50
N	0.6	20	5.9	4	19	50	6.4	2.8	3.3	7.1	5.1	3.6	3.4	1.3	0.8	3.7	4.4	20	14	14	1.2	2.8	3	4.8	1	3.6	6.1	5.5	0.2	3.9	1.6	15	4.9	1
P	11	20	0.4	7.1	20	20	20	2.2	20	6.7	20	5.3	20	20	4.9	20	20	20	20	18	20	20	9.6	18	20	20	5.7	17	2.5	17	4.6	0.3	20	8.7
Q	2.8	11	6.2	4.2	19	20	4.9	1.6	2	3.6	2.7	5.4	4.2	4	2.4	2.2	3.5	15	15	5.7	5.3	1.3	2.6	5.6	4	1.8	2	1.2	3.4	3.6	1.6	16	6.1	3.5
R	2.6	11	5.5	4.6	18	20	6.3	2.5	2.4	2.7	3	5.4	4.4	8.3	2.2	3.4	4.8	18	17	11	3.9	1.9	3.2	5.8	4.4	2.1	2	6.6	3.6	2.4	1.1	16	5.7	3.6
S	0.4	9.6	6.4	0.7	18	20	5.3	5.2	3.4	4.5	4.1	0.8	0.3	3.6	1	1.5	2.3	20	13	50	3.1	3.8	2.9	1	3.3	3.1	4.8	6.1	2.6	1.3	2.4	12	5.3	2.1
T	3.8	7.7	11	4.8	20	18	0.2	4.1	6.3	1.4	4.4	6.2	6.7	9.2	2.4	0.5	0.7	19	9.8	50	10	3.1	0.6	0.5	8.4	1.7	1.4	0.3	3.7	0.1	3	10	0.1	2.6
V	50	9.2	10	10	20	11	2.6	2.2	6.4	2.6	3.6	50	50	17	50	2	1.6	17	12	8.9	11	2	2.6	5.7	50	0.8	0.5	4.8	5.8	6.6	2.6	7.6	2.7	50
W	50	14	20	50	20	0.3	20	15	11	12	9.9	50	50	18	50	5.3	15	5	6.1	7.2	18	20	16	14	50	9.9	14	17	50	20	16	18	15	50
Y	50	8.7	18	50	19	4	20	8.7	5.4	6.4	2.8	50	50	20	50	3.1	14	0.2	1.4	0.6	20	3.6	11	50	50	50	9.6	50	50	20	8.9	15	9.8	50
	S	L	P	S	G	W	T	Q	L	T	K	A	S	D	D	T	T	Y	Y	Y	N	K	T	T	D	V	V	T	N	T	R	P	T	D

FIG. 5.

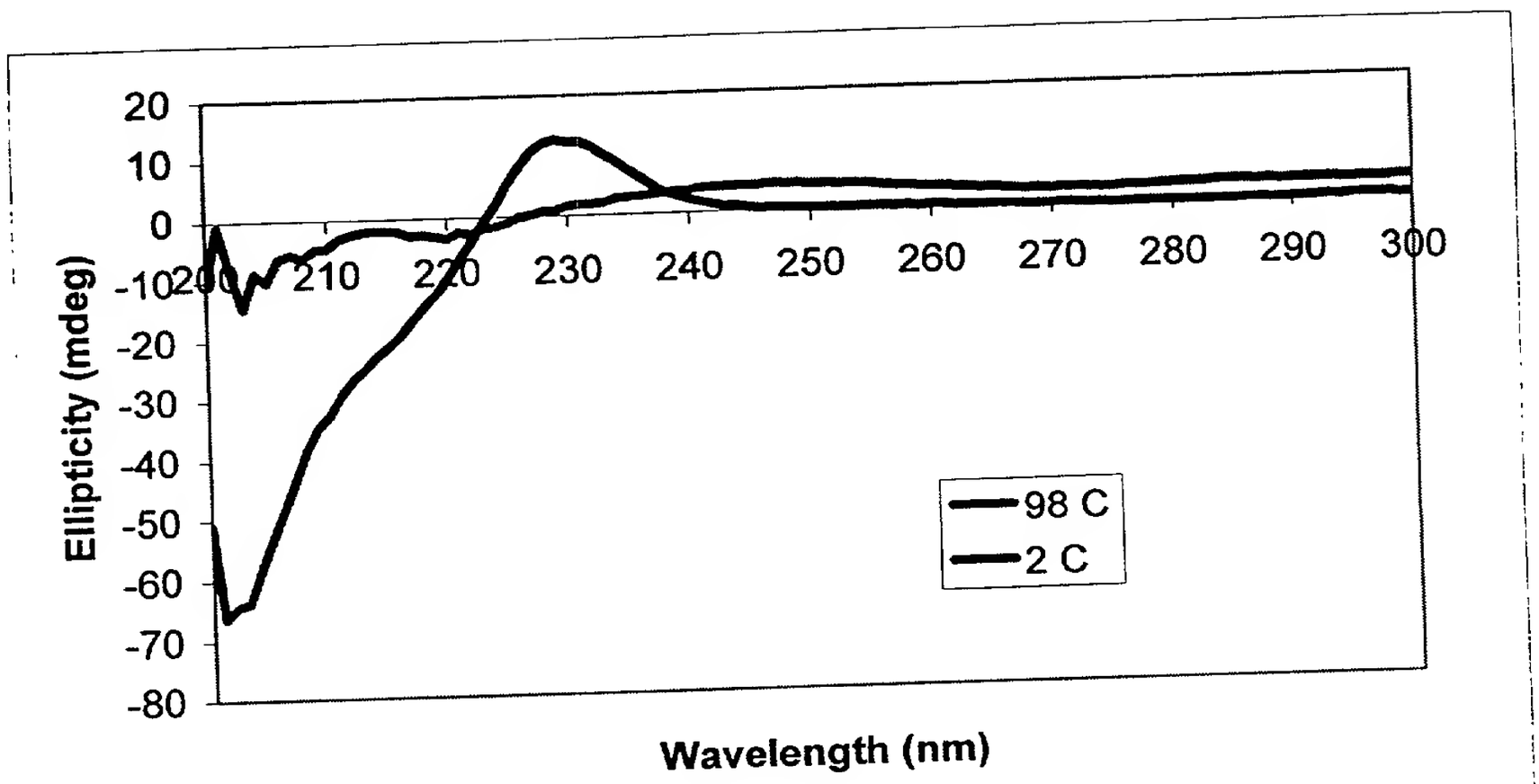
Ellipticity (mdeg)
20
10
0
-10
-20
-30
-40
-50
-60
-70
-80
200
210
220
230
240
250
260
270
280
290
300
98 C
2 C
Wavelength (nm)

FIG. 6.

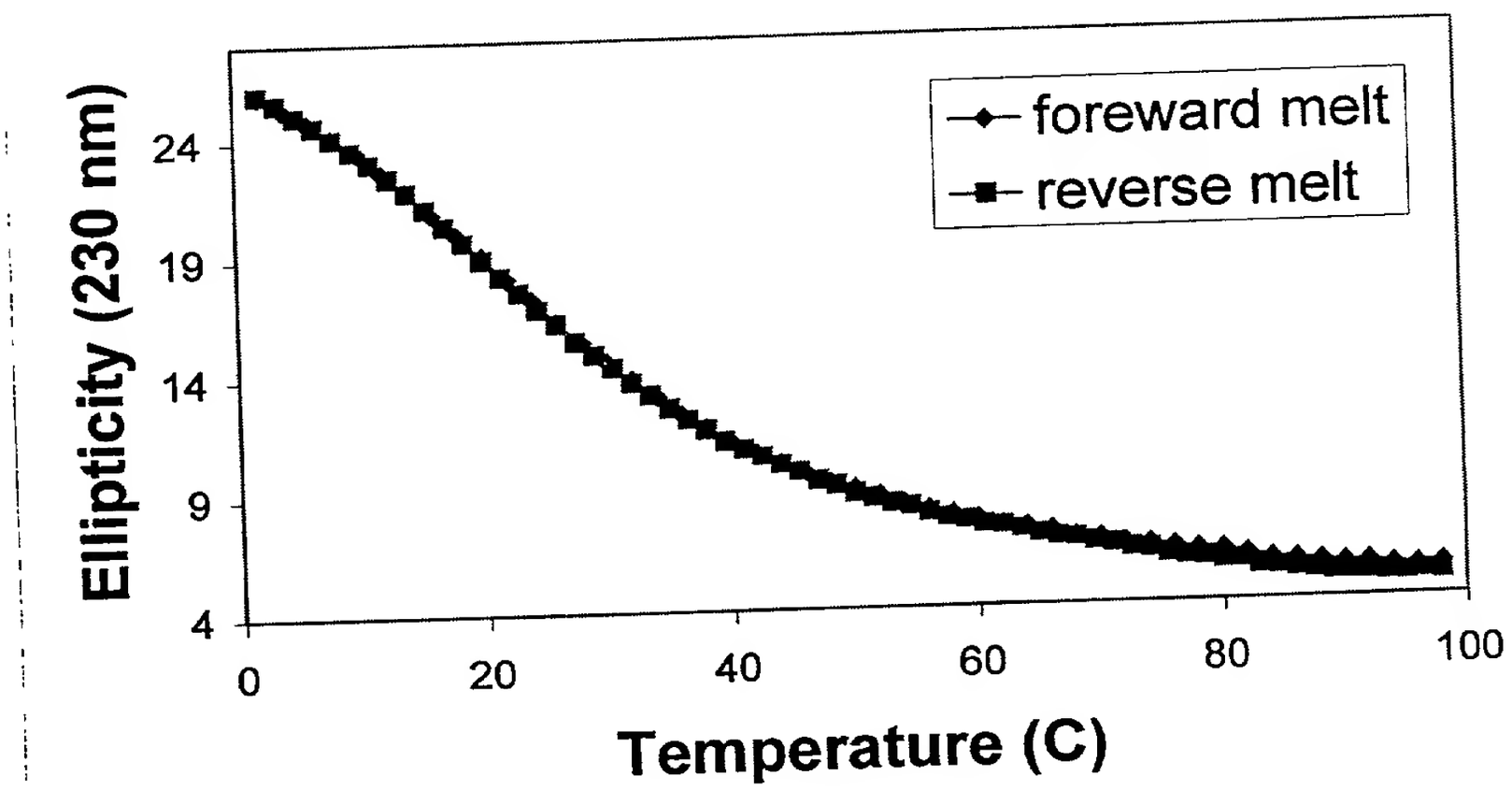
Ellipticity (230 nm)
24
19
14
9
4
0
20
40
60
80
100
foreward melt
reverse melt
Temperature (C)

FIG. 7.

FIG. 8A.

library design using free energy limits

complexity ~ 10^5

SLPSGWTQLTKASDDTTYYYNKTTDVVTNTRPTD													
N		K	K	SG	N			Q	SN	K		N	N
D		P			E							Q	
		V			S								
		E											
		R											

FIG. 8B.

complexity ~ 10^8

SLPSGWTQLTKASDDTTYYYNKTTDVVTNTRPTD													
N		KQK		SGNNSV		F		Q	SNTKQD			N	N
D		P			E			E		QT		Q	
E		V			S								
		E											
		R											
		N											

FIG. 8C.

library design using probability limit

complexity ~ 10^5

SLPSGWTQLTKASDDTTYYYNKTTDVVTNTRPTD													
N		K	K	SGNN	V	F		Q	SN	KQ		N	N
D													

FIG. 8.